

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Journal of Actuarial Practice 1993-2006

Finance Department


1993

Estimating the Effect of Statutory Changes on Insured Casualty Losses Using Generalized Indicator Variables

Ruy A. Cardoso

Automobile Insurers Bureau

Follow this and additional works at: <http://digitalcommons.unl.edu/joap>

 Part of the [Accounting Commons](#), [Business Administration, Management, and Operations Commons](#), [Corporate Finance Commons](#), [Finance and Financial Management Commons](#), [Insurance Commons](#), and the [Management Sciences and Quantitative Methods Commons](#)

Cardoso, Ruy A., "Estimating the Effect of Statutory Changes on Insured Casualty Losses Using Generalized Indicator Variables" (1993). *Journal of Actuarial Practice 1993-2006*. 172.
<http://digitalcommons.unl.edu/joap/172>

This Article is brought to you for free and open access by the Finance Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Journal of Actuarial Practice 1993-2006 by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Estimating the Effect of Statutory Changes on Insured Casualty Losses Using Generalized Indicator Variables

Ruy A. Cardoso*

Abstract

Techniques for estimating future insured losses in casualty insurance typically assume consistency in the insurance environment over time. Statutory changes, however, can create sharp discontinuities in the loss-generating process, complicating the estimation of those losses. Using indicator variables and dummy variables allows for quantification of the effect of such discontinuities. Three examples from private passenger automobile insurance are presented to illustrate how these variables can be used.

Key words and phrases: *dummy variables, linear regression, tort threshold, coverage stacking, coverage trigger, coverage limits*

1 Introduction

Estimation of future insured losses in casualty insurance often is based on an examination of the past patterns of those losses over time. Usually a linear or exponential relationship between losses and time is postulated as a starting point. Under this traditional actuarial approach, a further implicit assumption is that the insured losses are generated by an underlying process that changes smoothly. Statutory changes, however, can create discontinuities in the loss-generating process that must be accounted for properly in estimating future losses. This paper explains and illustrates a simple method of accounting for such discontinuities after they have occurred. Specifically, the method uses generalized forms of the linear regression variables known as *indicator (or dummy) variables*. Section 2 describes the most common actuarial method of estimating future losses in the absence of such discontinuities, while Section 3 provides

* Ruy Cardoso, FCAS, MAAA, currently is vice president and chief actuary of the Automobile Insurers Bureau of Massachusetts where he represents the Massachusetts automobile insurance industry in regulatory hearings concerning insurance rates. Mr. Cardoso previously has held consulting and insurance company positions whose primary focus was the analysis of casualty insurance loss reserves. He graduated from the Massachusetts Institute of Technology in 1983 with an S.B. in management science.

some background on indicator variables. Section 4 provides several specific examples of the generalized indicator variable approach using Massachusetts private passenger automobile insurance data, and Section 5 briefly summarizes the advantages and disadvantages of this approach.

2 Traditional Estimation of Future Losses

For simplicity, the discussion below assumes that the quantity of interest in the estimation procedures is the pure premium or the average insured loss per unit of insurance exposure. For private passenger automobile insurance, the unit of insurance exposure is generally a car-year, i.e., a single car insured for one year. The two most common models used to estimate future pure premiums assume either a linear or an exponential relationship between pure premiums (Y) and time (T), as shown in equations (1) and (2):

$$Y = a + bT \quad (1)$$

$$Y = a e^{bT}, \quad (2)$$

where a and b are constants (McClenahan, 1990). These two models often are based on economic indices rather than time and frequently include adjustments for autocorrelation (Cummins and Derrig, 1993). For simplicity in explaining the indicator variable approach, the remainder of this paper focuses on equation (1). Equation (2) sometimes is called *log-linear* because it can be transformed into equation (1) by taking logs. Once equation (2) is transformed, indicator variables also can be applied in a manner similar to that in equation (1). The interpretation of the quantities discussed below, however, would be different in the transformed case.

In practice, the time variable used in equation (1) is discrete, most often the accident year (the year in which the accident generating the loss occurred) associated with each loss. Further, the traditional method does not rely on individual losses. It works instead with aggregate pure premiums, in this case for each accident year. Thus, equation (1) simply says that pure premiums change by a constant dollar amount per year. Future pure premiums are estimated by assuming that the estimated annual change will continue into the future, although practicing actuaries often will modify the equation's results if its underlying assumptions are too strict.

It is not necessary to attribute the estimated pure premium change to specific causes, although blind application of the model

may lead to unreasonable results, especially if the random component of the loss-generating process is high. Pure premiums are not a direct (causal) function of the time variable; time is intended as a proxy for the many unspecified factors that determine pure premiums. This lack of causal explanation, however, is common to many possible methods of estimating future pure premiums. For example, one may use a Box-Jenkins¹ time series model (an approach widely used in non-actuarial settings) to relate the pure premium for a given accident year to pure premiums for past accident years and/or to past random errors, not to any underlying causal variables. The primary reason for using the time proxy is that, in practice, the number of available pure premium data points is usually too small to perform meaningful analyses of causal relationships (or, for that matter, Box-Jenkins analysis).

Whatever the underlying causal variables are, equation (1) implicitly assumes that they will behave smoothly over time. When there is a significant underlying change in the smoothness of the loss-generating process, the model is likely to produce poor estimates, making it necessary to deal with such discontinuities in some reasonable way. While the subjective adjustments frequently used in practice (for example, adjustment of data before the change to a postchange basis) may be appropriate in certain situations, the use of generalized indicator variables provides a more objective approach.

3 Background on Indicator Variables

An indicator random variable usually is defined with respect to the occurrence or non-occurrence of an event. Thus, if A is an event and $I(A)$ is the indicator random variable of A , then

$$I(A) = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise.} \end{cases}$$

In this paper, A is assumed to be an event (a change in the environment) that affects the pure premium. (See Miller and Wichern (1977) for a brief discussion of indicator variables in linear regression analysis.) Incorporation of indicator variables into equation (1) produces the model shown in equation (3):

¹ For a detailed description of the Box-Jenkins time series model and analysis, see Box and Jenkins (1970). For a brief introductory treatment, however, see Wheelwright and Makridakis (1985).

$$Y = a + bT + \sum_{i=1}^m c_i I_i(A_i) \quad (3)$$

where m is the number of indicator variables used, c_i , $i = 1, 2, \dots, m$ are constants, and I_i is the indicator variable for the i^{th} change. Table 1 illustrates the results of such a model when $m = 1$, $a = \$100$, $b = \$10$, $c_1 = \$5$, $I_1 = 0$ for $T \leq 3$, and $I_1 = 1$ for $T \geq 4$. Here A is the event $\{T \geq 4\}$.

TABLE 1
Hypothetical Pure Premium Model

T	Y	Change in Pure Premium	Indicator
1	\$110	\$10	0
2	\$120	\$10	0
3	\$130	\$10	0
4	\$145	\$15	1
5	\$155	\$10	1
6	\$165	\$10	1

Under equation (3), the indicator variable can be thought of as an on-off switch that reflects some change in the environment at and beyond $T = 4$. In a sense, a model using such a variable has one foot in the world of causal explanation.

It is not necessary for an indicator variable to be strictly zero-one, however. The terms *generalized indicator variables* and *dummy variables* are used interchangeably in this paper to reflect more general forms. Many changes in an environment are more analogous to a dimmer switch than to a simple on-off switch. That is, they occur gradually rather than all at once. McDowall, *et al.* (1980) describe the use of generalized indicator variables (or *intervention components* in their terminology) in Box-Jenkins time series analysis. The applications below will illustrate both the zero-one case and more general cases in the context of linear regressions against time, using statutory changes affecting private passenger automobile insurance as examples.

4 Specific Applications: Private Passenger Automobile Insurance

Permanent statutory changes in the insurance environment can have at least three effects on accident year pure premium data:

- a) **Single step**, reflecting a change that is completely effective in a specified accident year and all subsequent accident years.

- b) **Two step** (or, more generally, multiple step), reflecting a change that is partially effective in a specified accident year and completely effective in all subsequent accident years.
- c) **Infinite step**, reflecting a change that is partially effective in a specified accident year and increasingly effective in all subsequent accident years, but never completely effective.

The first effect can be modeled using the simple zero-one indicator variable, the second using an indicator variable that takes values between zero and one, and the third using an indicator variable that takes values greater than one.

4.1 The Single Step Case

An example of a single step statutory change in private passenger automobile insurance is a change in the *tort threshold*, the level of injuries that must be sustained before a person injured in an automobile accident can sue for pain and suffering damages. Certain states have no restrictions on the right to sue (i.e., there is no tort threshold), while those states where a no-fault system exists have either a qualitative threshold (usually referred to as a *verbal threshold*) or a *monetary threshold* (usually measured by medical costs). In the state of Massachusetts the current tort threshold is a monetary one. That is, the medical costs of the injuries sustained in an accident must exceed a fixed dollar amount before a suit for pain and suffering can be filed. On January 1, 1989 this threshold was raised from \$500 to \$2,000 for all accidents occurring on or after January 1, 1989. It follows that:

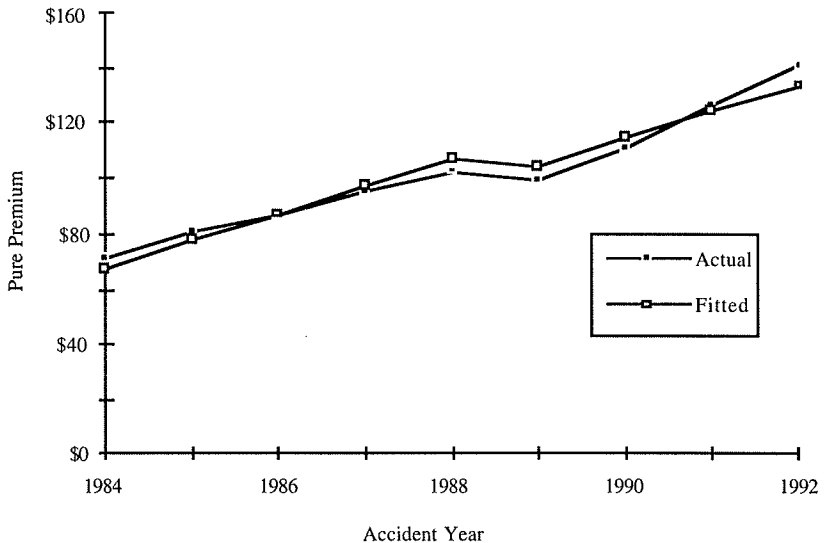
$$I = \begin{cases} 1 & \text{if accident year} \geq 1989 \\ 0 & \text{otherwise.} \end{cases}$$

Table 2 displays the accident year pure premiums for the bodily injury liability (BIL) coverage for the accident years 1984-1992. (Losses are limited to basic limits and developed to ultimate values.) Figure 1 displays the values in graphical form.

TABLE 2
Ultimate BIL Basic Limits Pure Premiums by Accident Year

Acc Year	T	BIL Pure Premium	Change in Pure Premium	Indicator
1984	1	\$71.01	NA	0
1985	2	\$80.28	\$9.27	0
1986	3	\$85.70	\$5.42	0
1987	4	\$95.35	\$9.65	0
1988	5	\$102.00	\$6.65	0
1989	6	\$99.01	(\$2.99)	1
1990	7	\$110.23	\$11.22	1
1991	8	\$125.64	\$15.41	1
1992	9	\$140.78	\$15.14	1

Figure 1
Ultimate BIL Basic Limits Pure Premiums by Accident Year



As both Table 2 and Figure 1 show, pure premiums rose steadily over the accident years 1984-1992 except in accident year 1989, the year in which the tort threshold was raised. As had been expected, raising the threshold reduces the pure premiums for the bodily injury liability coverage. An estimate of how much the pure premiums were reduced can be obtained using linear regression of the pure premiums against both the accident years and an indicator variable that is assigned the value of zero in accident years 1984-1988 and the value

of one in accident years 1989-1992. The regression results are shown in equation (4).

$$PP = 57.39 + 9.83 \times T - 12.19 \times I \quad (4)$$

where PP denotes pure premium, T denotes accident year (with 1984 considered year 1), and I denotes the indicator variable. Fitted values according to this equation also are displayed in Figure 1. The interpretation of equation (4) is that pure premiums are rising at \$9.83 per year and that the change in the tort threshold reduces pure premiums by \$12.19 from what they otherwise would have been (although the t -statistic for the coefficient of the indicator variable is not significantly different from zero under the usual significance levels). Future pure premiums in the presence of the higher tort threshold can be estimated using the above equation and holding the indicator variable at its postchange value of one. Naturally, the use of fewer data points will result in different estimates. This model's residuals indicate serial correlation of the errors, although the serial correlation might disappear if the infinite step model described in Section 4.3 were used. Analysis of residuals, however, is not a topic for this paper. Equation (4) simply serves to show how a zero-one indicator variable can be applied.

4.2 The Two-Step Case

Because private passenger automobile insurance policies are written throughout a given calendar year, the policy that covers an accident occurring in a particular accident year may have been written in that year or in the prior year. A change in the terms of the policy, therefore, will not affect all accidents occurring in a given year, only those covered by policies written after the change. In other words, a policy change will have only a partial effect on the accident year in which the change is made.

At the same time the tort threshold was raised in Massachusetts, another pair of statutory changes led to just this effect. A *stacking* provision (which determines whether policy limits from multiple policies in the same household can be combined) and a *trigger* provision (which determines the conditions under which coverage applies) were both modified in a way that was expected to reduce pure premiums. These modifications only applied to uninsured/underinsured motorists (UM/UIM) coverages, which pay for injuries in which a driver has insufficient bodily injury liability insurance (if any) to cover an insurance claim arising from an accident he or she caused.

Prior to January 1, 1989, households with more than one UM/UIM policy, under certain circumstances, could combine (stack) the limits of all of those policies to cover a single accident, in effect multiplying the limit on each policy by the number of policies in the household (if the limits were the same for each policy). This ability to combine limits was removed for policies written on or after January 1, 1989, reducing aggregate losses paid from what they otherwise would have been.

The change in the trigger provision works as follows. Losses paid under the UM/UIM coverages were unaffected by the limits of an at-fault driver's bodily injury liability insurance until January 1, 1989. Policies written on or after that date, however, only pay losses up to the difference in limits between the UM/UIM coverage, and the at-fault driver's bodily injury liability limits. (That is, an additional constraint must be satisfied before the coverage is triggered.)

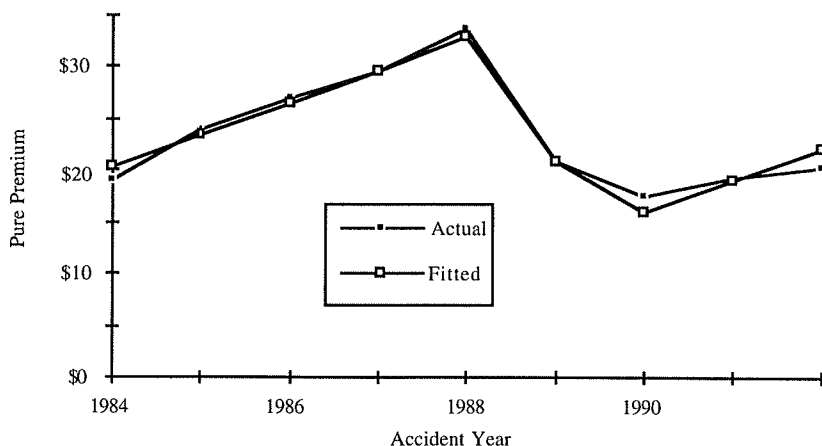
Because of the effective date of these changes, they were only partially effective in accident year 1989 but completely effective in all subsequent accident years. Based on the distribution of inception dates for policies written in Massachusetts, about 65 percent of the accidents occurring in accident year 1989 were covered by the modified policy.² Table 3 displays the accident year pure premiums for the UM/UIM coverages for the accident years 1984-1992. (Losses are limited to basic limits and developed to ultimate values.) Figure 2 displays the values in graphical form.

TABLE 3
Ultimate UM/UIM Basic Limits Pure Premiums by Accident Year

Acc Year	T	UM/UIM Pure Premium	Change in Pure Premium	Indicator
1984	1	\$18.91	NA	0
1985	2	\$23.83	\$4.92	0
1986	3	\$26.91	\$3.08	0
1987	4	\$29.40	\$2.49	0
1988	5	\$33.56	\$4.16	0
1989	6	\$20.91	(\$12.65)	0.65
1990	7	\$17.50	(\$3.41)	1
1991	8	\$19.27	\$1.77	1
1992	9	\$20.20	\$0.93	1

² The losses paid under the UM/UIM coverages also should have been affected by the change in the tort threshold, but to a far lesser degree than they were affected by the stacking and trigger changes.

Figure 2
Ultimate UM/UIM Basic Limits Pure Premiums by Accident Year



As both Table 3 and Figure 2 show, pure premiums rose steadily over the accident years 1984-1992 except in accident years 1989 and 1990, the two years over which the modified stacking and trigger provisions became effective. As expected, the two changes reduce the pure premiums for the UM/UIM coverages. An estimate of how much the pure premiums were reduced can be obtained using linear regression of the pure premiums against both the accident years and a generalized indicator variable (dummy variable) I that is assigned the value of zero in accident years 1984-1988, the value of 0.65 in accident year 1989, and the value of one in accident years 1990-1992, i.e.,

$$I = \begin{cases} 0 & \text{if } T = 1, 2, 3, 4, 5 \\ 0.65 & \text{if } T = 6 \\ 1 & \text{if } T \geq 7. \end{cases}$$

The regression results are shown in equation (5):

$$PP = 17.14 + 3.13 \times T - 23.17 \times I \quad (5)$$

where PP denotes pure premium, T denotes accident year (with 1984 considered year 1), and I denotes the dummy variable. Fitted values according to this equation also are displayed on Figure 2. The interpretation of equation (5) is that pure premiums are rising at \$3.13 per year and that the stacking and trigger modifications reduce pure premiums by \$23.17 from what they otherwise would have been. As the indicator variable for accident year 1989 is 0.65, the reduction in that year was not the full value of \$23.17, however; it was instead a partial value of $(0.65 \times \$23.17)$ or \$15.06. Future pure premiums under the modified stacking and trigger provisions can be estimated using the above equation and holding $I=1$. Again, the use of fewer data points will result in different estimates. Despite the two step nature of the discontinuity in this case, the functional form of the equation is the same as that of equation (4). Both are simply special cases of the general equation (3).

It is important to note, however, that the two step case described above also could be modeled using two zero-one indicator variables, the first changing to one in 1989 and the second changing to one in 1990. While the results of such a model would be similar to the results produced by equation (5) (due to the close fit), they would come at the cost of a degree of freedom and a less apparent model structure. It is easy to grasp the concept of a partial effect by seeing a generalized indicator variable with a value of 0.65, and it is clear in this instance that the 0.65 value has an objective basis rather than one that only pretends not to steal a degree of freedom.

4.3 The Infinite-Step Case

Certain changes in the insurance environment not only shift the relationship between pure premiums and time but also change the slope of the relationship. This type of effect can be modeled using two generalized indicator variables, the first the usual zero-one type and the second comprising a series of infinitely increasing values. The particular change in Massachusetts that can be modeled this way occurred at the same time as the change in the tort threshold and was effective for all accidents occurring in accident year 1989 and subsequent years (despite contrary policy language). Specifically, the coverage limit of the personal injury protection (PIP) coverage increased from \$2,000 to \$8,000 on January 1, 1989. This coverage pays for injuries regardless of fault and therefore also is known as *no-fault coverage*.

Because many of the claims paid under the PIP coverage reached the \$2,000 limit in the years before the limit was increased, claim

cost inflation only could affect a subset of all claims. Increasing the limit to \$8,000, however, allows those claims previously constrained by the limit to reflect the effects of claim cost inflation, in turn allowing the aggregate pure premiums for the PIP coverage to reflect inflation more completely and thus increase more quickly (i.e., with a greater slope). While it is possible that increasing the tort threshold also may have a slope-changing effect on BIL coverage (see Section 4.1), the PIP limit change serves as a much clearer illustration.

If we denote the slope of the pure premium line under the \$2,000 limit as b , the size of the discontinuity created by the statutory change as c_1 , and the slope of the pure premium line under the \$8,000 limit as c_2 (where c_2 is expected to be greater than b), then pure premiums over time can be modeled as follows:

$$PP = \begin{cases} a + bT & \text{for } T \leq 5 \\ a + bT + c_1 & \text{for } T = 6 \\ a + 6b + c_1 + c_2(T-6) & \text{for } T \geq 7. \end{cases} \quad (6A)$$

While this is a natural way to model the PIP pure premiums over time, equation (6A) does not fit into the general equation (3). In order to transform equation (6A) into a specific instance of equation (3), it is necessary to redefine c_2 as the difference between the post-1989 slope and the pre-1989 slope (where the difference is expected to be positive) and adopt the following pair of generalized indicator variables:

$$I_1 = \begin{cases} 0 & \text{if } T \leq 5 \\ 1 & \text{if } T \geq 6 \end{cases}$$

and

$$I_2 = \begin{cases} 0 & \text{if } T \leq 6 \\ (T-6) & \text{if } T \geq 7. \end{cases}$$

Equation (6A) can be recast as:

$$PP = \begin{cases} a + bT & \text{for } T \leq 5 \\ a + bT + c_1 & \text{for } T = 6 \\ a + bT + c_1 + c_2(T-6) & \text{for } T \geq 7, \end{cases} \quad (6B)$$

with equation (6B) being a specific instance of equation (3):

$$PP = a + bT + c_1I_1 + c_2I_2. \quad (6C)$$

Table 4 below displays the accident year pure premiums for the PIP coverage for the accident years 1984 to 1992 (again developed to ultimate values); Figure 3 displays the values in graphical form.

TABLE 4
Ultimate PIP Pure Premiums by Accident Year

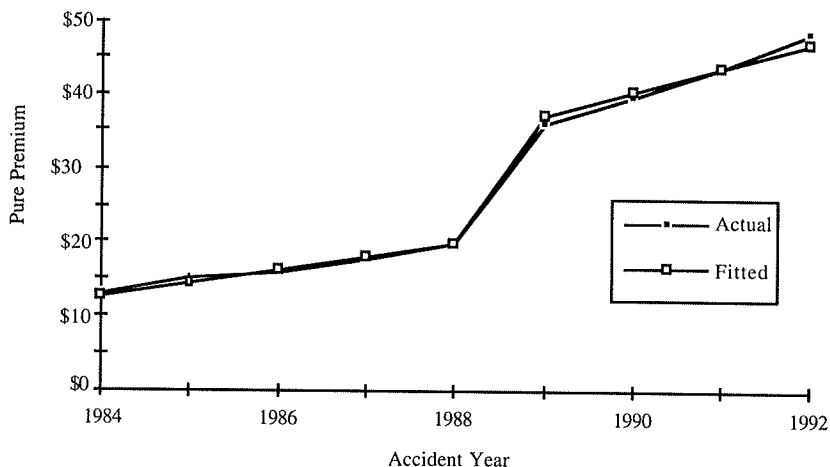
Acc Year	T	PIP Pure Premium	Change in Pure Premium	Indicator #1	Indicator #2
1984	1	\$12.98	NA	0	0
1985	2	\$14.97	\$1.99	0	0
1986	3	\$15.92	\$0.95	0	0
1987	4	\$17.61	\$1.69	0	0
1988	5	\$19.63	\$2.02	0	0
1989	6	\$36.03	\$16.40	1	0
1990	7	\$39.81	\$3.78	1	1
1991	8	\$43.39	\$3.58	1	2
1992	9	\$48.33	\$4.94	1	3

As both Table 4 and Figure 3 show, pure premiums rose steadily over accident years 1984 through 1988, jumped sharply at accident year 1989, and rose more steeply over accident years 1990 through 1992 (as expected). An estimate of how much the pure premium line was shifted and steepened because of the change in limit can be obtained using linear regression of the pure premiums against both the accident years and the two indicator variables displayed in Table 4 above. The regression results are shown in equation (7).

$$PP = 11.44 + 1.59 \times T + 14.81 \times I_1 + 2.45 \times I_2 \quad (7)$$

where PP again denotes pure premium, T denotes accident year, and I_1 and I_2 denote the indicator variables. Fitted values according to this equation are displayed on Figure 3. The interpretation of equation (7) is that pure premiums were rising at \$1.59 per year, increased \$14.81 as a result of the change in the PIP coverage limit (because $I_1=1$ in 1989), and now are rising at \$4.04 per year, where \$4.04 equals the

Figure 3
Ultimate PIP Pure Premiums By Accident Year



prechange slope of \$1.59 plus the postchange increment of \$2.45 (the coefficient of I_2). Future pure premiums under the \$8,000 PIP coverage limit can be estimated by using the above equation, holding I_1 constant at its value of one and moving I_2 up one for each year beyond accident year 1989.

Relative to the single step and two step cases, this case has cost another degree of freedom. But in this situation an additional quantity is being estimated, specifically the postchange slope, making the cost an appropriate one to pay. Further, the model structure is reasonably apparent. While other approaches could be used to model the infinite step case, the one used here strikes the best balance between clarity and degrees of freedom.

5 Summary of the Approach

As illustrated above, generalized indicator variables can be used to model a variety of different time series discontinuities in private passenger automobile insurance. While the examples above have

been restricted to permanent statutory changes, the approach can be extended easily to temporary changes as well as to other lines of insurance. This flexibility is a key advantage of the approach, as is its ability to let the data speak for themselves. Alternative approaches, such as adjusting all data to a postdiscontinuity basis, can work in the single step case above, but such an alternative is likely to be more subjective than the generalized indicator variable approach.

On the other hand, a too-complicated set of indicator variables could be used to mask the occasional tendency to force a preordained conclusion. Further, the use of multiple indicator variables easily could lead to overfitting, especially in the common situation where only a small number of data points is available. Such pitfalls should not blind the actuary to the usefulness of the generalized indicator variable approach. As with any model-building exercise, the value of indicator variables as a tool will rise with the care taken in using them.

References

- Box, G.E.P., and Jenkins, G.M. *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day, 1970.
- Cummins, J.D., and Derrig, R.A. "Fuzzy Trends in Property-Liability Insurance Claim Costs." *Journal of Risk and Insurance* 60, 3 (September 1993): 429-465.
- McClenahan, C.L. "Ratemaking," Chapter 2 of *Foundations of Casualty Actuarial Science*. New York, New York: Casualty Actuarial Society, 1990.
- McDowall, D., et al. *Interrupted Time Series Analysis*. Newbury Park, California: Sage Publications, Inc., 1980.
- Miller, R.B. and Wichern, D.W. *Intermediate Business Statistics: Analysis of Variance, Regression, and Time Series*. New York, New York: Holt, Rinehart and Winston, 1977.
- Wheelwright, S.C. and Makridakis, S. *Forecasting Methods for Management*, 4th Edition. New York, New York: John Wiley & Sons, Inc. 1985.

Ruy A. Cardoso
Automobile Insurers Bureau of Massachusetts
101 Arch Street
Boston, Massachusetts 02110